

A unified approach to source location via rooted network augmentation

Zeev Nutov

The Open University of Israel nutov@openu.ac.il

Abstract. In **Source Location (SL)** problems the goal is to select a minimum cost source set S such that the connectivity from S to any node v is at least the demand d_v of v . In **Network Augmentation (NA)** problems we are given a graph $G = (V, E)$ and an edge-set F on V , edge-costs on F or node-costs on V , and connectivity requirements $\{r_{sv} : sv \in D\}$ on a set D of “demand edges”. The goal is to compute a minimum cost edge-set $I \subseteq F$, such that in the graph $G + I$, the connectivity of each $sv \in D$ is at least r_{sv} . In **Rooted NA** D is a star with center s , and in **a -Based NA** F is a star with center a . We suggest a notion of q -connectivity, where every node v has a capacity $q_v \geq 1$ that represents the resistance of v to failures. We show that a large class of **SL** problems, including the variants that appear in the literature, is a particular case of q -connectivity **s -Based Rooted NA** problems. We use this to derive some approximation algorithms for **SL** from those for **NA**, as well as to derive some new results for **SL** problems. Some of our results are as follows.

- (i) We show that the $(1 + \ln |D|)$ -approximation algorithm of [12] for **a -Based NA** with edge-costs has ratio $1 + \ln \min\{\sum_{uv \in D} r_{uv}, |D| \cdot p_{\max}\}$ for the q -connectivity version with node-costs, where p_{\max} is the maximum number of parallel edges in F . In particular, this implies by a simple proof the $(1 + \ln \sum_{v \in V} d_v)$ -approximation algorithms of Sakashita, Makino, and Fujishige [19] for three **SL** variants. Moreover, we will show that this is so whenever the connectivity function is submodular.
- (ii) For undirected **a -Based NA** we give an approximation algorithm with ratio $\min\{k, p_{\max} \ln k\} \cdot O(\ln k)$ for node-costs and $O(\ln^2 k)$ for edge-costs; the latter improves the ratio $\tilde{O}(k)$ of Fukunaga [5].

1 Introduction

1.1 Problems considered and relations between them

In this paper we suggest a unifying approach to handle **Source Location** problems via **Rooted Network Augmentation** problems. In **Source Location** problems the goal is to select a minimum cost source set S such that the connectivity from S to any other node v is at least the demand d_v of v . In **Network Augmentation** problems the goal is to augment a given graph by a minimum-cost edge-set such that the new graph satisfies prescribed connectivity requirements. Formally, the generic versions of these problems are as follows. Given a function $w = \{w_u : u \in U\}$ on

a groundset U and $U' \subseteq U$, let $w(U') = \sum_{u \in U'} w_u$. If w is a cost function on U and I is an edge-set on U , then the cost (or the node-costs) $w(I)$ of I is defined to be the cost of the set of the endnodes of I . Also, let $w_{\max} = \max_{u \in U} w_u$ denote the maximum value of w .

Source Location (SL)

Instance: A graph $G = (V, E)$ with node-costs $c = \{c_v : v \in V\}$, connectivity demands $d = \{d_v : v \in V\}$, and connectivity function $\psi : 2^V \times V \rightarrow \mathbb{Z}_+$.

Objective: Find a minimum cost source node set $S \subseteq V$ such that $\psi(S, v) \geq d_v$ for every $v \in V$.

Network Augmentation (NA)

Instance: A graph $G = (V, E)$ and an edge-set F on V , a cost function c on F or on V , connectivity requirements $r = \{r_{sv} : sv \in D\}$ on a set D of demand edges on V , and a family $\{f_{sv} : 2^F \rightarrow \mathbb{Z}_+ : sv \in D\}$ of connectivity functions.

Objective: Find a minimum-cost edge-set $I \subseteq F$ such that $f_{sv}(I) \geq r_{sv}$ for every $sv \in D$.

In NA problems, typical connectivity functions are as follows.

- *Edge-connectivity* $\lambda_G(s, v)$ is the maximum number of pairwise edge disjoint sv -paths in G .
- *Node-connectivity* $\kappa_G(s, v)$ is the maximum number of pairwise internally disjoint sv -paths in G .
- *Q -connectivity* $\lambda_G^Q(s, v)$ for given $Q \subseteq V$, is the maximum number of sv paths no two of which have an edge or an internal node in Q in common. Note that Q -connectivity reduces to edge-connectivity if $Q = \emptyset$, and to node-connectivity if $Q = V$; namely, $\lambda_G^\emptyset(s, v) = \lambda_G(s, v)$ and $\lambda_G^V(s, v) = \kappa_G(s, v)$.

The corresponding versions of NA are edge-connectivity NA when $f_{sv}(I) = \lambda_{G+I}(s, v)$, node-connectivity NA when $f_{sv}(I) = \kappa_{G+I}(s, v)$, and Q -connectivity NA when $f_{sv}(I) = \lambda_{G+I}^Q(s, v)$.

Most papers that considered SL problems defined (S, v) -edge-connectivity $\lambda_G(S, v)$ as the maximum number of pairwise edge-disjoint (S, v) -paths. On the other hand, several definitions were used for (S, v) -node-connectivity $\kappa(S, v)$; in most of these definition $\kappa(S, v) = \infty$ if $v \in S$. Here we suggest a definition that captures previous definitions as particular cases.

Note that $\lambda_G^Q(s, v)$ is the maximum sv -flow value, where edges and nodes in Q have unit capacities, while the capacity of the nodes in $V \setminus Q$ is ∞ . Hence the Q -connectivity function $\lambda_G^Q(s, v)$ is the max-flow/min-cut value function in G with node-capacities in $\{1, \infty\}$ and unit edge capacities. In Definition 1 below, this is generalized as follows. Every node v has capacity q_v (the resistance of v to failures), and an amount p_v of flow-supply that v can deliver to any other node (including itself) if v is chosen to be in the source set S .

Definition 1. Let $G = (V, E)$ be a graph with node-capacities $\{q_v : v \in V\}$. For $S \subseteq V$ and $v \in V$ the (S, v) - q -connectivity $\lambda_G^q(S, v)$ is the maximum flow value from S to v (every edge has capacity 1). Equivalently, $\lambda_G^q(S, v)$ is the minimum capacity of a cut $C \subseteq E \cup V \setminus \{v\}$ such that $G \setminus C$ has no (S, v) -path. Given flow-supplies $\{p_v : v \in S\}$, the (S, v) -(p, q)-connectivity $\lambda_G^{p,q}(S, v)$ is the maximum (s, v) -flow value $\lambda_{G'}^q(s, v)$ in the network G' obtained by adding to G a new node s and connecting it to every $u \in S$ with p_u edges; hence, $\lambda_G^{p,q}(S, v) = p_v + \lambda_G^q(S, v)$ if $v \in S$, and if $p_v \geq q_v$ for every $v \in S$ then $\lambda_G^{p,q}(S, v) = \lambda_G^q(s, v)$.

Note that $\lambda_G(S, v) = \lambda_G^{p,q}(S, v)$ for $p_v = \infty$ and $q_v = \infty$ for every $v \in V$, and that $\lambda_G^Q(s, v) = \lambda_G^q(s, v)$ for $q_v = 1$ if $v \in Q$ and $q_v = \infty$ otherwise. Now we mention some node-connectivity functions ψ that appear in the literature, and show that they are particular case of the (p, q) -connectivity function $\lambda_G^{p,q}$ with $p_v, q_v \in \{1, \infty\}$ and $q_v \leq p_v$.

1. $\kappa(S, v)$ is the maximum number of (S, v) -paths no two of which have a common node in $V \setminus (S \cup v)$ if $v \notin S$, and $\kappa(S, v) = \infty$ otherwise; equivalently, $\kappa(S, v)$ is the minimum size $|C|$ of a cut $C \subseteq E \cup V \setminus (S \cup \{v\})$ such that $G \setminus C$ has no (S, v) -path.

For directed graphs, κ -SL is equivalent to λ -SL by the following (approximation ratio preserving) standard reduction: replace every $v \in V$ by two nodes v^{in}, v^{out} connected by an edge $v^{in}v^{out}$, and replace every edge $uv \in E$ by an edge $u^{out}v^{in}$; each node v^{out} inherits the cost and the demand of v , while v^{in} has cost ∞ and demand 0. It is not hard to verify that S is a feasible solution to the original instance with connectivity function κ if, and only if, $S^{out} = \{v^{out} : v \in S\}$ is a feasible solution to the obtained instance with the edge-connectivity function λ . For undirected graphs, we do not see that κ is a particular case of $\lambda_G^{p,q}$, but we are also not aware on any work on *undirected* κ -SL problems.

2. $\hat{\kappa}(S, v)$ is the maximum number of (S, v) -paths no two of which have a common node in $V \setminus \{v\}$ if $v \notin S$, and $\hat{\kappa}(S, v) = \infty$ otherwise; equivalently, $\hat{\kappa}(S, v)$ is the minimum size $|C|$ of a cut $C \subseteq E \cup V \setminus (\{v\})$ such that $G \setminus C$ has no (S, v) -path.

Then $\hat{\kappa}(S, v) = \lambda_G^{p,q}(S, v)$ for $p_v = \infty$ and $q_v = 1$ for every $v \in V$.

3. $\kappa'(S, v) = \hat{\kappa}(S, v)$ if $v \notin S$ and $\kappa'(S, v) = 1 + \hat{\kappa}(S \setminus \{v\}, v)$ if $v \in S$; equivalently, $\kappa'(S, v)$ is the minimum size $|C|$ of a cut $C \subseteq E \cup V \setminus (\{v\})$ such that $G \setminus C$ has no $(S \setminus \{v\}, v)$ -path.

Then $\kappa'(S, v) = \lambda_G^{p,q}(S, v)$ for $p_v = 1$ and $q_v = 1$ for every $v \in V$.

Given an instance of SL or of NA, let k denote the maximum demand $d_{\max} = \max_{v \in V} d_v$ or the maximum requirement $r_{\max} = \max_{uv \in D} r_{uv}$. Note that in SL problems with $\psi(S, v) = \lambda_G^{p,q}(S, v)$, we may always assume that $p_v \leq k$, and it is also reasonable to assume that $1 \leq q_v \leq p_v$ (as in the above versions). We consider a relation between such (p, q) -connectivity versions of SL and q -connectivity versions of NA, that formally are defined as follows.

Survivable Source Location (SSL)

This is SL with connectivity function $\psi(S, v) = \lambda_G^{p,q}(S, v)$ and $1 \leq q_v \leq p_v \leq k$ for all $v \in V$.

Survivable Network Augmentation (SNA)

This is NA with connectivity functions $f_{sv}(I) = \lambda_{G+I}^q(s, v)$.

Now we define several particular important cases of NA.

Rooted NA

A particular case of NA when D is a star, whose center we denote by s .

 a -Based NA

A particular case of NA when F is a star centered at a .

Connectivity Augmentation is a particular case of Q -connectivity SNA, when any edge can be added to G by a cost of 1.

a -Based Connectivity Augmentation is a particular case of Q -connectivity SNA, when any edge leaving a can be added to G by a cost of 1.

The a -Based Connectivity Augmentation problem was defined in [12], where it was also shown to admit a $(\ln r(D) + 1)$ -approximation algorithm. The study of this problem in [12] is motivated by the following observation.

Observation 1 ([12]) *If directed a -Based Connectivity Augmentation admits approximation ratio ρ , then directed Connectivity Augmentation admits the following approximation ratios: 2ρ if $Q \neq V$, and $2k\rho$ if $Q = V$.*

For an edge-set/graph J let $\delta_J(X)$ denote the set of edges in J from X to $V \setminus X$. This paper is motivated by a recent paper of Fukunaga [5], that defined the connectivity function κ' , and observed that κ' -SSL is equivalent to the particular case of s -Based Rooted SNA with *edge-costs* and with $|\delta_G(s)| = c(s) = 0$. Here we further observe the following.

Observation 2 *For both directed and undirected graphs, SSL is equivalent to s -Based Rooted SNA with *node-costs* and with $|\delta_G(s)| = c_s = 0$.*

Proof. The reduction is essentially the one in Definition 1. Given an instance of SSL construct an instance of s -Based Rooted SNA as follows: add to G a new node s of cost 0, and for every $v \in V$ set $r_{sv} = d_v$ and put a_v edges from s to v into F . Conversely, given an instance of s -Based Rooted SNA with node-costs and $|\delta_G(s)| = c_s = 0$, construct an instance of SSL as follows. Remove s from G , and for every $v \in V$ set p_v is the number of edges in F from s to v and $d_v = r_{sv}$. In both directions, it is not hard to see that S is a solution to the SSL instance, if, and only if, the edge set I of all edges in F from s to S is a solution to the s -Based Rooted SNA instance, and clearly I and S have the same node-cost. \square

Note that in the above reduction, we have the following.

- The case of uniform demands (namely, $d_v = k$ for all $v \in V$) in SSL corresponds to the case of rooted uniform requirements (namely $r_{sv} = k$ for all $v \in V \setminus \{s\}$) in s -Based Rooted SNA.
- The case of unit costs (namely, $c_v = 1$ for all $v \in V$) in SSL corresponds to the version of s -Based Connectivity Augmentation with node-costs, when we can pick p_v edges from s to v by a cost of 1.

Let us illustrate the usefulness of above relation between the two problems. Directed Rooted SNA with edge-costs and uniform requirements $r_{sv} = k$ for all $v \in V$ can be solved in polynomial time [4]; this easily implies that also undirected Rooted s -Based SNA with edge-costs can be solved in polynomial time. Thus the same holds for κ' -SSL, since κ' -SSL is a particular case of s -Based Rooted SNA with edge-costs. Frank [3] showed that *directed s -Based Connectivity Augmentation* with $\delta_G(s) = \emptyset$ is NP-hard. Using a slight modification of his reduction we can show that the problem is in fact **Set-Cover** hard to approximate, and thus is $\Omega(\log n)$ -hard to approximate. Given an instance of **Set-Cover**, where a family A of sets needs to cover a set B of elements, construct the corresponding directed bipartite graph $G' = (A \cup B, E')$, by putting an edge from every set to each element it contains. The graph $G = (V, E)$ is obtained from G' by adding M copies of B , connecting A to each copy in the same way as to B , and adding a new node s . Let $F = \{sv : v \in V\}$, $c(e) = 1$ for every $e \in F$, and $r_{sv} = 0$ if $v \in A$ and $r_{sv} = 1$ otherwise. It is easy to see that if $I \subseteq F$ is a feasible solution to the obtained Rooted s -Based SNA instance, then either I corresponds to a feasible solution to the **Set-Cover** instance, or $|I| \geq M$. The $\Omega(\log n)$ -hardness follows for M large enough, say $|M| = (|A| + |B|)^2$, and $|A| = |B|$. Since for $k = 1$ all connectivity functions of SSL are equivalent, we get $\Omega(\log n)$ hardness for directed SSL with $k = 1$ and unit costs. We summarize this as follows.

Corollary 3. *Directed SSL for $k = 1$ and unit costs is $\Omega(\log n)$ -hard to approximate. Directed/undirected κ' -SSL with uniform demands can be solved in polynomial time.*

1.2 Previous work

The previously best known approximation ratios and hardness of approximation results for SL problems with connectivity functions $\lambda, \kappa, \hat{\kappa}, \kappa'$, including the ones in Corollary 3, are summarized in Table 1; see also a survey in [15].

Some additional results are as follows. Ishii, Fujita, and Nagamochi [7,8] showed that undirected SSL with $\hat{\kappa}$ can be solved in polynomial time for $k \leq 3$, but is NP-hard if there exists a vertex $v \in V$ with $d(v) \geq 4$. Barasz, Becker, and Frank [2] gave a strongly polynomial time algorithm for edge-connectivity directed SSL with uniform demands. Several generalizations of source location problems can be found in [20,9,14].

$c \ \& \ d$	$\lambda \ (p, q \equiv \infty)$		κ	
	<i>Undirected</i>	<i>Directed</i>	<i>Undirected</i>	<i>Directed</i>
GC & GD	$\Theta(\ln d(V))$ [19]	$\Theta(\ln d(V))$ [19]	$\Theta(\ln d(V))$ [19]	$\Theta(\ln d(V))$ [19]
GC & UD	in P [1]	$O(\ln d(V))$ [19]	$O(\ln d(V))$ [19]	$O(\ln d(V))$ [19]
UC & GD	in P [1]	$O(\ln d(V))$ [19]	$O(\ln d(V))$ [19]	$O(\ln d(V))$ [19]
UC & UD	in P [22]	in P [10]	$O(\ln d(V))$ [19]	$O(\ln d(V))$ [19]
	$\hat{\kappa} \ (p \equiv \infty, q \equiv 1)$		$\kappa' \ (p, q \equiv 1)$	
	<i>Undirected</i>	<i>Directed</i>	<i>Undirected</i>	<i>Directed</i>
GC & GD	$\Theta(\ln d(V))$ [19] $O(k \ln k)$ [5]	$\Theta(\ln d(V))$ [19]	$O(\ln d(V))$ [5] $O(k \ln k)$ [5]	$O(\ln d(V))$ [5]
GC & UD	in P [16]	in P [16]	in P	in P
UC & GD	$O(\ln d(V))$ [19] $O(k)$ [6]	$O(\ln d(V))$ [19]	$O(\ln d(V))$ [5]	$O(\ln d(V))$ [5] $\Omega(\log n)$
UC & UD	in P [16]	in P [16]	in P	in P

Table 1. Previous approximation ratios and hardness of approximation thresholds for SL problems. GC and UC stand for general and uniform costs, GD and UD stand for general and uniform demands, respectively.

1.3 Our results

Observations 1 and 2 motivate the study of the α -Based SNA problem. Interestingly, the algorithms of [19] for SL with connectivity functions $\lambda, \kappa, \hat{\kappa}$, and the algorithm of [12] for α -Based Connectivity Augmentation both use the same method of reducing the problem to a submodular covering problem. Now we see that is not a coincidence, since by Observations 1 and 2, both problems are particular cases of the α -Based SNA problem. Furthermore, we will show by a simple proof that such a reduction is possible whenever the connectivity function is submodular and non-decreasing. A set function f on U is submodular if $f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$ for all $X, Y \subseteq U$, and f is non-decreasing if $f(X) \leq f(Y)$ for all $X \subseteq Y \subseteq U$. Our first result is obtained by observing that the algorithm of [12] for α -Based Connectivity Augmentation extends to α -Based NA with arbitrary submodular non-decreasing connectivity function, for both edge-costs and node-costs, as follows. Recall that for an α -Based SNA instance, p_{\max} denotes the maximum number parallel edges in F , and that $p_{\max} \leq k$.

Theorem 1. *If for a directed α -Based SNA instance each function $f_{uv}(I)$ is submodular and non-decreasing, then the problem admits an approximation algorithm with ratio $1 + \ln \min\{r(D), \alpha\}$, where $\alpha = \max_{e \in F} \sum_{uv \in D} f_{uv}(\{e\})$ in the case of edge costs, and $\alpha = \max_{z \in V} \sum_{uv \in D} f_{uv}(\delta_F(z))$ in the case of node costs.*

In Section 3 we prove that for α -Based SNA instances, the set-function on F defined by $f_{uv}(I) = \lambda_{G+I}^q(u, v)$ is submodular and non-decreasing, and that $\max_{e \in F} \sum_{uv \in D} f_{uv}(\{e\}) \leq |D|$ and $\max_{z \in V} \sum_{uv \in D} f_{uv}(\delta_F(z)) \leq |D| \cdot p_{\max}$, where p_{\max} is the maximum number of parallel edges in F .

Theorem 2. *Directed α -Based SNA admits approximation ratios $1 + \ln |D|$ for edge-costs and $1 + \ln \min\{r(D), |D| \cdot p_{\max}\}$ for node-costs. Thus directed SSL admits approximation ratio $1 + \ln \min\{r(D), |D| \cdot p_{\max}\}$.*

We note that for both SSL and SNA, an approximation ratio ρ for directed graphs implies approximation ratio 2ρ for undirected graphs. Usually, undirected connectivity problems are easier to approximate than the directed ones. Directed SSL is already Set-Cover hard even for $k = 1$ and unit costs, but for undirected SSL with $k = 1$ (and even with $k = 2$) it is not hard to obtain a polynomial time algorithm. Hence a natural question is whether undirected SSL admits an approximation ratio that depend on k only. As was mentioned, Fukunaga [5] observed that Q -connectivity SSL with $Q = V$ is equivalent to s -Based Rooted SNA with edge-costs. Since undirected Q -connectivity Rooted SNA with edge-costs admits ratio $O(k \ln k)$ [17], so is Q -connectivity SSL with $Q = V$. However, Connectivity Augmentation admits ratio $O(\ln^2 k)$ for rooted requirements [18], and a natural question to ask is whether the algorithm of [18] extends to s -Based Rooted SNA.

Theorem 3. *Undirected α -Based SNA admits the following approximation ratios, where $H(k) = \sum_{i=1}^k 1/i$ denotes the k th Harmonic number.*

- (i) *For edge-costs, $\sum_{\ell=1}^k \frac{H(\Delta_\ell)}{k-\ell+1} \leq H(k) \cdot H(\Delta_\ell) = O(\ln^2 k)$, where $\Delta_\ell = (4\ell - 3)^2$; s -Based Rooted SNA admits ratio $\sum_{\ell=1}^k \frac{H(2\ell-1)}{k-\ell+1} \leq H(k) \cdot H(2k-1)$.*
- (ii) *For node-costs, $\sum_{\ell=1}^k H(\Delta_\ell) \min\left\{\frac{p_{\max}}{k-\ell+1}, 1\right\} \leq \min\{k, p_{\max}H(k)\} \cdot H(\Delta_k)$; s -Based Rooted SNA, and thus also undirected SSL, admits ratio $\sum_{\ell=1}^k H(2\ell-1) \min\left\{\frac{p_{\max}}{k-\ell+1}, 1\right\} \leq \min\{k, p_{\max} \cdot H(k)\} \cdot H(2k-1)$.*

Part (ii) implies that undirected SSL admits approximation ratio that depends on k only; this ratio is better than the one implied by Theorem 2 for $k = o\left(\frac{\ln n}{\ln \ln n}\right)$. Furthermore, it improves the ratio $O(k \ln k)$ of Fukunaga [5] for κ' -SSL to $O(\ln^2 k)$, since we have $p_{\max} = 1$ in this case.

Theorems 1, 2, and 3, are proved in Sections 2, 3, and 4, respectively.

2 Proof of Theorem 1

We use a result due to Wolsey [23] about a performance of a greedy algorithm for submodular covering problems. A generic covering problem is as follows.

Covering Problem

Instance: A groundset U with costs $\{c_u : u \in U\}$ and an integer function g on 2^U given by an evaluation oracle.

Objective: Find $A \subseteq U$ of minimum cost such that $g(A) = g(U)$.

The *Greedy Algorithm* starts with $A = \emptyset$ and as long as $g(A) < g(U)$ repeatedly adds to A an element $u \in U \setminus A$ such that $\frac{g(A \cup \{u\}) - g(A)}{c_u}$ is maximum.

In [23] it is proved that if p is non-decreasing and submodular, and if $p(\emptyset) = 0$, then the greedy algorithm has approximation ratio $1 + \ln \max_{u \in U} g(u)$.

We start with the case of edge-costs. Then the function g is defined in the same way as in [12]. We have $U = F$ and for $I \subseteq F$

$$g(I) = r(D) - \sum_{uv \in D} \max\{r(u, v) - f_{uv}(I), 0\}.$$

It is not hard to verify that I is a feasible solution to **NA** instance if, and only if, $g(I) = g(F) = r(D)$, that $g(\emptyset) = 0$, and that g is non-decreasing. Also, we have

$$\max_{e \in F} g(\{e\}) = r(D) - \max_{e \in F} \sum_{uv \in D} \max\{r(u, v) - f_{uv}(\{e\}), 0\} \leq \min\{r(D), \alpha\}.$$

We show that g is submodular. A set function h is supermodular if $-h$ is submodular. It is known (c.f. [21]) that if h is supermodular, then $r + h$ is supermodular for any constant r , and that $\max\{h, 0\}$ is supermodular. Thus the function $h_{uv}(I) = \max\{r(u, v) - f_{uv}(I), 0\}$ is supermodular. As a sum of supermodular functions is also supermodular, we obtain that g is submodular.

Now let us consider node-costs. For $S \subseteq V$ let F_S denote the set of edges in F from a to S , and let $f'_{uv}(S) = f_{uv}(F_S)$. We have $U = V$ and for $S \subseteq V$ let

$$g'(S) = r(D) - \sum_{uv \in D} \max\{r(u, v) - f'_{uv}(S), 0\}.$$

As in edge-costs case, it is not hard to verify that S is a feasible solution to an **NA** instance if, and only if, $g'(S) = g'(V) = r(D)$, that $g'(\emptyset) = 0$, and that g' is non-decreasing. Also, we have

$$\max_{z \in V} g'(\{z\}) = r(D) - \max_{z \in V} \sum_{uv \in D} \max\{r(u, v) - f'_{uv}(\{z\}), 0\} \leq \min\{r(D), \alpha\}.$$

We show that g' is submodular. We claim that the submodularity of $f(I)$ implies that $f'(S)$ is submodular. This is not true in general, but holds if F is a star, and hence for *a*-Based **NA** instances. More precisely, it is not hard to verify the following statement, that finishes the proof of Theorem 1.

Lemma 1. *Let (V, F) be a graph and let f be a submodular set function on F . If F is a star, then the set function on V defined by $f'(S) = f(F_S)$ is also submodular.* \square

3 Proof of Theorem 2

All graphs in this section are assumed to be directed. Theorem 2 will follow from Theorem 1 and the following Lemma, whose parts were implicitly proved in [12].

Lemma 2. *For any directed *a*-Based SNA instance, for any $s, v \in V$, the set-function $f_{sv}(I) = \lambda_{G+I}^q(s, v)$ on F is submodular and non-decreasing. Furthermore, $\max_{e \in F} \sum_{sv \in D} f_{sv}(\{e\}) \leq |D|$ and $\max_{z \in V} \sum_{sv \in D} f_{sv}(\delta_F(z)) \leq |D| \cdot p_{\max}$.*

In the rest of this section we prove Lemma 2. Let $s, v \in V$. It is easy to see that $f_{sv}(I) = \lambda_{G+I}^q(u, v)$ is non-decreasing. Also, it is clear that $f_{sv}(I) \leq |I|$ for any $I \subseteq F$, and thus $f_{sv}(\{e\}) \leq 1$ for any $e \in F$ and $f_{sv}(\delta_F(z)) \leq |\delta_F(z)| \leq p_{\max}$. Thus it remains to prove that $f(I) = f_{sv}(I) = \lambda_{G+I}^q(s, v)$ is submodular. For that, we will use the following known characterization of submodularity, c.f. [21]: *A set-function f on F is submodular if, and only if*

$$f(I_0 \cup \{e\}) + f(I_0 \cup \{e'\}) \geq f(I_0) + f(I_0 \cup \{e, e'\}) \quad \forall I_0 \subseteq F, e, e' \in F \setminus I_0$$

Let us fix $I_0 \subseteq F$. Revising our notation to $G \leftarrow G + I_0$, $F \leftarrow F \setminus I_0$, and denoting $h(I) = f(I_0 \cup I) - f(I_0)$, we get that f is submodular if, and only if

$$h(\{e\}) + h(\{e'\}) \geq h(\{e, e'\}) \quad \forall e, e' \in F.$$

In our setting, $h(I) = \lambda_{G+I}^q(s, v) - \lambda_G^q(s, v)$ is the increase in the (s, v) - q -connectivity as a result of adding I to G . Thus $0 \leq h(I) \leq |I|$ for any $I \subseteq F$, so $0 \leq h(\{e, e'\}) \leq 2$. If $h(\{e, e'\}) = 0$, then we are done; if $h(\{e, e'\}) = 1$, then we need to show that $h(\{e\}) = 1$ or $h(\{e'\}) = 1$; and if $h(\{e, e'\}) = 2$, then we need to show that $h(\{e\}) = 1$ and $h(\{e'\}) = 1$. We prove the following general statement, that implies the above.

Lemma 3. *Let $G = (V, E)$ be a directed graph with node capacities $\{q_v : v \in V\}$, let I be a set of edges on V disjoint to E , let $s, t \in V$, and let $h = \lambda_{G+I}^q(s, t) - \lambda_G^q(s, t)$. Then there is $J \subseteq I$ of size $|J| \geq h$ such that $\lambda_{G+J}^q(s, t) = \lambda_G^q(s, t) + 1$ for every $e \in J$.*

Proof. Since we consider directed graphs, it is sufficient to prove the lemma for the case of edge-connectivity. For that, apply the following standard reduction that eliminates node capacities: replace every $v \in V \setminus \{s, t\}$ by two nodes v^{in}, v^{out} connected by q_v parallel edges from v^{in} to v^{out} and replace every edge $uv \in E \cup F$ by an edge from u^{out} to v^{in} . Hence we will prove the lemma for the edge connectivity function λ . Let us say that $S \subseteq V$ is *tight* if $s \in S$, $v \notin S$, and $|\delta_G(S)| = \lambda_G(s, v)$. Let \mathcal{F} be the family of tight sets. By Menger's Theorem \mathcal{F} is non-empty. It is known that \mathcal{F} is a ring family, namely, the intersection of all the sets in \mathcal{F} is nonempty, and if $X, Y \in \mathcal{F}$ then $X \cap Y, X \cup Y \in \mathcal{F}$. Then \mathcal{F} has a unique inclusion-minimal set S_{\min} and a unique inclusion-maximal set S_{\max} . Let $J = \{uv \in I : u \in S_{\min}, v \in V \setminus S_{\max}\}$ be the set of edges in I that go from S_{\min} to $V \setminus S_{\max}$. By Menger's Theorem, $|J| \geq h$, and $\lambda_{G+J}(s, t) = \lambda_G(s, t) + 1$ for any $e \in J$. The statement follows. \square

The proof of Theorem 2 is complete.

4 Proof of Theorem 3

Here we prove Theorem 3. All graphs in this and the next section are assumed to be undirected, unless stated otherwise. We start by considering the edge-costs case, and then will show that it implies the node-costs case by reductions.

Definition 2. An ordered pair $\hat{X} = (X, X^+)$ of subsets of a groundset V is called a *biset* if $X \subseteq X^+$; X is the inner part and X^+ is the outer part of \hat{X} , and $\Gamma(\hat{X}) = X^+ \setminus X$ is the boundary of \hat{X} . An edge e covers a biset \hat{X} if it has one endnode in X and the other in $V \setminus X^+$. For a biset \hat{X} and an edge-set/graph J let $\delta_J(\hat{X})$ denote the set of edges in J covering \hat{X} .

Given an instance of **SNA** and a biset \hat{X} on V , let the requirement of \hat{X} be $r(\hat{X}) = \max\{r_{uv} : uv \in \delta_D(\hat{X})\}$ if $\delta_D(\hat{X}) \neq \emptyset$ and $r(\hat{X}) = 0$ otherwise. By the q -connectivity version of Menger's Theorem (c.f. [11]), $I \subseteq F$ is a feasible solution to an **SNA** instance if, and only if, $|\delta_I(\hat{X})| \geq h(\hat{X})$ for every bisets \hat{X} on V , where h is a biset-function defined by

$$h(\hat{X}) = \max\{r(\hat{X}) - (q(\Gamma(\hat{X})) + |\delta_G(\hat{X})|), 0\} \quad (1)$$

Let \mathcal{P}_h denote the polytope of “fractional edge-covers” of h , namely,

$$\mathcal{P}_h = \left\{ x \in \mathbb{R}^F : x(\delta_F(\hat{Y})) \geq h(\hat{Y}) \ \forall \text{ biset } \hat{Y} \text{ on } V, \ 0 \leq x_e \leq 1 \ \forall e \in F \right\}.$$

Let $\tau^*(h)$ denote the optimal value of a standard LP-relaxation for edge covering h by a minimum cost edge set, namely, $\tau^*(h) = \min \left\{ \sum_{e \in F} c_e x_e : x \in \mathcal{P}_h \right\}$.

As an intermediate problem, we consider **SNA** instances when we seek to increase the connectivity by 1 for every $uv \in D$, namely, when $r_{uv} = \lambda_{G+I}^q(u, v) + 1$ for all $uv \in D$.

D-SNA (the edge-costs version)

Instance: A graph $G = (V, E)$ and an edge set F on V , node-capacities $\{q_v : v \in V\}$, a cost function c on F , and a set D of demand edges on V .

Objective: Find a minimum cost edge-set $I \subseteq E$ such that $\lambda_{G+I}^q(u, v) \geq \lambda_G^q(u, v) + 1$ for all $uv \in D$.

Given a *D-SNA* instance, we say that a biset \hat{X} is *tight* if $h(\hat{X}) = 1$, where h is defined by (1). *D-SNA* is equivalent to the problem of finding a minimum cost edge-cover of the biset family $\mathcal{F} = \{\hat{X} : h(\hat{X}) = 1\}$ of tight bisets. Thus the following generic problem includes the *D-SNA* problem.

Biset-Family Edge-Cover

Instance: A graph (V, F) with edge-costs and a biset family \mathcal{F} on V .

Objective: Find a minimum cost \mathcal{F} -cover $I \subseteq F$.

For a biset-family \mathcal{F} let $\tau^*(\mathcal{F})$ denote the optimal value of a standard LP-relaxation for edge covering \mathcal{F} by a minimum cost edge set, namely, $\tau^*(\mathcal{F}) = \tau^*(h)$ for $h(\hat{X}) = 1$ if $\hat{X} \in \mathcal{F}$ and $h(\hat{X}) = 0$ otherwise.

Proposition 1. Suppose that *a-Based D-SNA with edge-costs* admits a polynomial time algorithm that computes a solution of cost at most $\rho(k)\tau^*(\mathcal{F})$, where \mathcal{F} is the family of tight bisets. Then *a-Based SNA* admits a polynomial time algorithm that computes a solution I such that:

- For edge-costs, $c(I) \leq \tau^*(h) \cdot \sum_{\ell=1}^k \frac{\rho(\ell)}{k-\ell+1}$, where h is defined by (1).
- For node-costs, $c(I) \leq \text{opt} \cdot \sum_{\ell=1}^k \rho(\ell) \cdot \min \left\{ \frac{p_{\max}}{k-\ell+1}, 1 \right\}$.

Proof. We start with the edge-costs case. Consider the following sequential algorithm. Start with $I = \emptyset$. At iteration $\ell = 1, \dots, k$, add to I and remove from F an edge-set $I_\ell \subseteq F$ that increases by 1 the q -connectivity of $G + I$ on the set of demand edges $D_\ell = \{sv : \lambda_{G+I}^q(s, v) = r(s, v) - k + \ell - 1, sv \in D\}$, by covering the corresponding biset-family \mathcal{F}_ℓ using the ρ -approximation algorithm. After iteration ℓ , we have $\lambda_{G+I}^q(s, v) \geq r(s, v) - k + \ell$ for all $sv \in D$. Consequently, after k iterations $\lambda_{G+I}^q(s, v) \geq r(s, v)$ holds for all $sv \in D$, thus the computed solution is feasible. The approximation ratio follows from the following two observations.

- (i) $c(I_\ell) \leq \rho(\ell) \cdot \tau^*(\mathcal{F}_\ell)$. This is so since $\lambda(s, v) \leq \ell - 1$ for every $sv \in D_\ell$, hence the maximum requirement at iteration ℓ is at most ℓ .
- (ii) $\tau^*(\mathcal{F}_\ell) \leq \frac{\tau^*(h)}{k-\ell+1}$. To see this, note that if $\hat{Y} \in \mathcal{F}_\ell$ and $x \in \mathcal{P}_h$ then $x(\delta(\hat{Y})) \geq k - \ell + 1$, by Menger's Theorem. Thus $x/(k - \ell + 1)$ is a feasible solution for the LP-relaxation for edge-covering \mathcal{F}_ℓ , of value $c \cdot x/(k - \ell + 1)$.

Consequently, $c(I) = \sum_{\ell=1}^k c(I_\ell) \leq \sum_{\ell=1}^k \rho(\ell) \cdot \frac{\tau^*(h)}{k-\ell+1} = \tau^*(h) \cdot \sum_{\ell=1}^k \frac{\rho(\ell)}{k-\ell+1}$.

Now let us consider the case of node-costs. Then we convert node-costs into edge-costs by assigning to every edge $e = av$ the cost $c'(e) = c(v)$. Let opt' denote the optimal solution value of the edge-costs instance obtained. Clearly, $\text{opt} \leq \text{opt}' \leq p_{\max} \cdot \text{opt}$. Note that any inclusion minimal solution to an a -Based D -SNA instance has no parallel edges. This implies that $c(I_\ell) \leq \rho(\ell) \cdot \text{opt}$ and that $c(I_\ell) = c'(I_\ell)$. The latter implies $c(I_\ell) = c'(I_\ell) \leq \rho(\ell) \cdot \frac{\text{opt}'}{k-\ell+1} \leq \rho(\ell) \cdot \text{opt}' \cdot \frac{p_{\max}}{k-\ell+1}$, and the statement for the node-costs case follows. \square

In the next section we will prove the following theorem, that together with Proposition 1 finishes the proof of Theorem 3.

Theorem 4. *For edge-costs, undirected a -Based D -SNA admits a polynomial time algorithm that computes a feasible solution I of cost $c(I) \leq H(\Delta_k) \cdot \tau^*(\mathcal{F})$, and $c(I) \leq H(2k - 1) \cdot \tau^*(\mathcal{F})$ if D is a star with center $s = a$, where \mathcal{F} is the family of tight bisets.*

5 Proof of Theorem 4

Recall that D -SNA is equivalent to Biset-Family Edge-Cover with \mathcal{F} being the family of tight bisets; in the case of rooted requirements, it is sufficient to cover the biset-family $\mathcal{F}^s = \{\hat{X} \in \mathcal{F} : s \in V \setminus X^+\}$. Biset-families arising from SNA instances have some special properties, that are summarized in the following definitions.

Definition 3. *The intersection and the union of two bisets \hat{X}, \hat{Y} is defined by $\hat{X} \cap \hat{Y} = (X \cap Y, X^+ \cap Y^+)$ and $\hat{X} \cup \hat{Y} = (X \cup Y, X^+ \cup Y^+)$. The biset $\hat{X} \setminus \hat{Y}$ is defined by $\hat{X} \setminus \hat{Y} = (X \setminus Y^+, X^+ \setminus Y)$. We write $\hat{X} \subseteq \hat{Y}$ and say that \hat{Y} contains \hat{X} if $X \subseteq Y$ and $X^+ \subseteq Y^+$. Let $\mathcal{C}_{\mathcal{F}}$ denote the inclusion-minimal bisets in \mathcal{F} .*

Definition 4. Two bisets \hat{X}, \hat{Y} covered by an edge-set D are D -independent if for any $xx', yy' \in D$ such that xx' covers \hat{X} and yy' covers \hat{Y} , $\{x, x'\} \cap \Gamma(\hat{Y}) \neq \emptyset$ or $\{y, y'\} \cap \Gamma(\hat{X}) \neq \emptyset$; otherwise, \hat{X}, \hat{Y} are D -dependent. We say that a biset family \mathcal{F} is D -uncrossable if D covers \mathcal{F} and if for any D -dependent $\hat{X}, \hat{Y} \in \mathcal{F}$ the following holds:

$$\hat{X} \cap \hat{Y}, \hat{X} \cup \hat{Y} \in \mathcal{F} \text{ or } \hat{X} \setminus \hat{Y}, \hat{Y} \setminus \hat{X} \in \mathcal{F}. \quad (2)$$

Similarly, given a set $T \subseteq V$ of terminals, we say that \hat{X}, \hat{Y} are T -independent if $X \cap T \subseteq \Gamma(\hat{Y})$ or if $Y \cap T \subseteq \Gamma(\hat{X})$, and \hat{X}, \hat{Y} are T -dependent otherwise. We say that \mathcal{F} is T -uncrossable if T covers the set-family of the inner parts of \mathcal{F} , and if (2) holds for any T -dependent $\hat{X}, \hat{Y} \in \mathcal{F}$.

A biset-family \mathcal{F} is symmetric if $\hat{X} \in \mathcal{F}$ implies $(V \setminus X^+, V \setminus X) \in \mathcal{F}$. Clearly, the family of tight bisets is symmetric. We will use the the following fundamental statement, that was implicitly proved in [18].

Lemma 4 ([18]). *The family \mathcal{F} of tight bisets is D -uncrossable, and if D is a star with center s and leaf-set T then $\{\hat{X} \in \mathcal{F} : s \notin X^+\}$ is T -uncrossable.*

For a biset-family \mathcal{C} let $\gamma_{\mathcal{C}} = \max\{|\Gamma(\hat{C})| : \hat{C} \in \mathcal{C}\}$. Note that if \mathcal{F} is the family of tight bisets then $\gamma_{\mathcal{F}} \leq k - 1$. Given an instance of **Biset-Family Edge-Cover**, we will assume that the family \mathcal{C} of the inclusion members of \mathcal{F} can be computed in polynomial time. We note that for \mathcal{F} being the family of tight sets, this step can be implemented in polynomial time, c.f. [18]. Under this assumption, we prove the following generalization of Theorem 4.

Theorem 5. *For edge/node-costs, a -Based Biset-Family Edge-Cover admits a polynomial time algorithm that computes a cover I of \mathcal{F} such that:*

- (i) $c(I) \leq H \left((4\gamma_{\mathcal{C}} + 1)^2 \right) \cdot \tau^*(\mathcal{F})$ if \mathcal{F} is symmetric and D -uncrossable.
- (ii) $c(I) \leq H(2\gamma_{\mathcal{C}} + 1) \cdot \tau^*(\mathcal{F})$ if \mathcal{F} is T -uncrossable and $a \in V \setminus X^+$ for all $\hat{X} \in \mathcal{F}$.

In the rest of this section we prove Theorem 5.

Definition 5. A node set $U \subseteq V$ is a \mathcal{C} -transversal of a hypergraph (set-family) \mathcal{C} on V if U intersects every set in \mathcal{C} ; if \mathcal{C} is a biset-family then U should intersect the inner part of every member of \mathcal{C} . Given costs $\{c_v : v \in V\}$, let $t^*(\mathcal{C})$ denote the minimum value of a fractional \mathcal{C} -transversal, namely:

$$t^*(\mathcal{C}) = \min \left\{ \sum_{v \in V} c_v x_v : x(C) \geq 1 \quad \forall \hat{C} \in \mathcal{C}, \quad x(v) \geq 0 \quad \forall v \in V \right\}.$$

In [18], the following is proved.

Theorem 6 ([18]). *Let \mathcal{F} be a biset-family and let \mathcal{C} be the family of the inclusion members of \mathcal{F} . Then the maximum degree in the hypergraph $\{\hat{C} : \hat{C} \in \mathcal{C}\}$ is at most:*

- (i) $(4\gamma_{\mathcal{C}} + 1)^2$ if \mathcal{F} is D -uncrossable.
- (ii) $2\gamma_{\mathcal{C}} + 1$ if \mathcal{F} is T -uncrossable.

Given a hypergraph (V, \mathcal{C}) with node-costs, the greedy algorithm computes in polynomial time a \mathcal{C} -transversal $U \subseteq V$ of cost $c(U) \leq H(\Delta(\mathcal{C}))t^*(\mathcal{C})$, where $\Delta(\mathcal{C})$ is the maximum degree of the hypergraph (c.f. [13]).

Lemma 5. *If an edge-set I covers a biset-family \mathcal{F} then the set of endnodes of I is a transversal of \mathcal{F} .*

Lemma 6. *Let \mathcal{F} be a biset family on V and I a star with center a on a transversal $U \subseteq V$ of \mathcal{F} . Then I covers \mathcal{F} in each one of the following cases.*

- (i) \mathcal{F} is symmetric and $a \notin \Gamma(\hat{X})$ for all $\hat{X} \in \mathcal{F}$.
- (ii) $a \in V \setminus X^+$ for all $\hat{X} \in \mathcal{F}$.

Proof. Let $\hat{X} \in \mathcal{F}$. Then $a \in X$ or $a \in V \setminus X^+$. If $a \in V \setminus X^+$, then since U is a transversal of \mathcal{C} , there is $u \in U \cap X$. If $a \in X$, then if \mathcal{F} is symmetric, then there $u \in U \cap (V \setminus X^+)$. In both cases, there is an edge $au \in I$, and this edge covers \hat{X} . \square

The algorithm as in Theorem 5, for both edge-costs and node-costs is as follows, where in the case of node-costs we may assume that the cost of a is zero.

1. For every $v \in V \setminus \{a\}$, let e_v be the minimum-cost edge incident to v , and in the case of edge-costs define node-costs $c_v = \min_{e \in \delta_F(v)} c_e$ if $\delta_F(v) \neq \emptyset$, and $c_v = \infty$ otherwise.
2. Let \mathcal{C} be the family of the inclusion members of \mathcal{F} . With node-costs $\{c_v : v \in V\}$, compute a transversal U of \mathcal{C} of cost $c(U) \leq H(\Delta(\mathcal{C}))t^*(\mathcal{C})$.
3. Return $I = \{e_v : v \in U\}$.

The solution computed is feasible by Lemma 6. The approximation ratio follows from Theorem 6 and Lemma 5.

The proof of Theorem 4 is complete.

Acknowledgment I thank Takuro Fukunaga for many useful comments and for his help in filling Table 1.

References

1. K. Arata, S. Iwata, K. Makino, and S. Fujishige. Locating sources to meet flow demands in undirected networks. *J. Algorithms*, 42:54–68, 2002.
2. M. Barasz, J. Becker, and A. Frank. An algorithm for source location in directed graphs. *Operations Research Letters*, 33:221–230, 2005.
3. A. Frank. Augmenting graphs to meet edge-connectivity requirements. *SIAM Journal on Discrete Mathematics*, 5(1):25–53, 1992.
4. A. Frank and E. Tardos. An application of submodular flows. *Linear Algebra and its Applications*, 114/115:329–348, 1989.

5. T. Fukunaga. Approximating minimum cost source location problems with local vertex-connectivity demands. In *TAMC*, pages 428–439, 2011.
6. T. Ishii. Greedy approximation for source location problem with vertex-connectivity requirements in undirected graphs. *Journal of Discrete Algorithms*, 7:570–578, 2009.
7. T. Ishii, H. Fujita, and H. Nagamochi. Source location problem with local 3-vertex-connectivity requirements. In *Proc. 3rd Hungarian-Japanese Symposium on Discrete Mathematics and Its Applications*, pages 368–377, 2003.
8. T. Ishii, H. Fujita, and H. Nagamochi. Minimum cost source location problem with local 3-vertex-connectivity requirements. *Theoretical Computer Science*, 372:81–93, 2007.
9. T. Ishii and K. Makino. Posi-modular systems with modulatone requirements under permutation constraints. *Discrete Mathematics, Algorithms and Applications*, 2(1):61–76, 2010.
10. H. Ito, K. Makino, K. Arata, S. Honami, Y. Itatsu, and S. Fujishige. Source location problem with flow requirements in directed networks. *Optimization Methods and Software*, 18(4):427–435, 2003.
11. G. Kortsarz and Z. Nutov. Approximating minimum cost connectivity problems, Ch. 58 in *Approximation Algorithms and Metaheuristics*, T. F. Gonzales ed. PWS, 2007.
12. G. Kortsarz and Z. Nutov. Tight approximation algorithm for connectivity augmentation problems. *J. Computer and System Sciences*, 74(5):662–670, 2008.
13. L. Lovász. On the ratio of optimal integral and fractional covers. *Discrete Mathematics*, 13:383–390, 1975.
14. S. Mamada, T. Uno, K. Makino, and S. Fujishige. An $O(n \log^2 n)$ algorithm for the optimal sink location problem in dynamic tree networks. *Discrete Applied Mathematics*, 154(16):2387–2401, 2006.
15. H. Nagamochi and T. Ibaraki. *Algorithmic Aspects of Graph Connectivity, Chapter 9*. Cambridge University Press, 2008.
16. H. Nagamochi, T. Ishii, and H. Ito. Minimum cost source location problem with vertex-connectivity requirements in digraphs. *Information Processing Letters*, 80:287–294, 2001.
17. Z. Nutov. Approximating minimum cost connectivity problems via uncrossable bifamilies. To appear in *Transactions on Algorithms*. Preliminary version in *FOCS* 2009, 417–426.
18. Z. Nutov. Approximating node-connectivity augmentation problems. *Algorithmica*, 63(1-2):398–410, 2012.
19. M. Sakashita, K. Makino, and S. Fujishige. Minimum cost source location problems with flow requirements. *Algorithmica*, 50:555–583, 2008.
20. M. Sakashita, K. Makino, H. Nagamochi, and S. Fujishige. Minimum transversals in posi-modular systems. *SIAM J. on Discrete Mathematics*, 23:858–871, 2009.
21. A. Schrijver. *Combinatorial Optimization Polyhedra and Efficiency*. Springer, 2004.
22. H. Tamura, M. Sengoku, S. Shinoda, and T. Abe. Location problems on undirected flow networks. *IEICE Trans.*, E73:1989–1993, 1990.
23. L. A. Wolsey. An analysis of the greedy algorithm for the submodular set covering problem. *Combinatorica*, 2:385–393, 1982.